AD_____

GRANT NUMBER: DAMD17-94-J-4414

TITLE: cDNA Fingerprinting of Breast Cancer Tumor Cells

PRINCIPAL INVESTIGATOR: Cassandra L. Smith, Ph.D.

CONTRACTING ORGANIZATION: Boston University
Boston, Massachusetts 02215

REPORT DATE: October 1995

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1995	Annual 30 Sep 94 - 29 Sep 95

4. TITLE AND SUBTITLE
cDNA Fingerprinting of Breast Cancer Tumor Cells

5. FUNDING NUMBERS
DAMD17-94-J-4414

6. AUTHOR(S)
Cassandra L. Smith , Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston University
Boston, Massachusetts 02215

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In the time period covered by this report we have carried out three sets of preliminary studies. Pulsed field gel electrophoresis has been used to examine possible amplified regions on chromosome20 implicated in breast cancer to provide target areas of the genome for further study. Various methods for producing dense arrays have been developed and evaluated. Elaborations on these methods have been designed that promise to produce high quality cDNA profiling. A new method for profiling DNA changes has been developed that uses different genomic gel fragments in a single gel lane. This method appears to be extremely promising and its use in the search for DNA changes important in breast cancer will be a major focus of future studies.

14. SUBJECT TERMS
breast cancer, cDNA, PCR, chromosome 20, pfg

15. NUMBER OF PAGES
13

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

X Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

X In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

X In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

X In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

___ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

C L Smith 11/29/95

PI - Signature Date

Progress Report

Grant # DAMD17-94-J-4414
Title: "cDNA Fingerprinting of Breast Cancer Cells"
Principle Investigator: Cassandra L. Smith

Table of Contents

Progress Report
Grant # DAMD17-94-J-4414
Title: cDNA Fingerprinting of Breast Cancer Cells
Principle Investigator: Cassandra L. Smith

Introduction.

The objectives of this proposal are to identify important changes at the DNA or RNA levels associated with specific breast cancer characteristics. These changes may occur through point mutations, or larger DNA rearrangements or amplification. A variety of different analytical methods is needed to deal with the wide range situations. The initial project plans relied heavily on two techniques. We have carried out a number of pilot studies on these techniques. In addition we have developed a new technique that appears particularly promising, a technique for profiling DNA sequence mixtures. This new technique will now form a major thrust of the project. The work currently ongoing is focused in three areas. (1) Genomic DNA analysis. Pulsed field gel electrophoresis fragments have been tested and found suitable as probes for fluorescence in situ hybridization of candidate breast cancer regions (collaboration with Joe Gray). A set of five cell lines with different breast cancer phenotypes has been obtained and high molecular weight DNA from these lines has been prepared and it is being screened for possible expansions of chromosome 20 that might be breast cancer associated. (2) cDNA analysis. A set of test cDNA clones has been selected and obtained from Incyte Pharmaceuticals. This set will be used to optimize and validate comparative cDNA hybridization methods. Several different ways for preparing stable reusable arrays of DNAs on surfaces have been evaluated and compared. The best of these will be chosen for large scale preparation of cDNA arrays. Preliminary studies on detection of fluorescent labeled nucleic acids in arrays have been carried out. The results of these studies allowed us to formulate the technical specifications of our detection system, and the components of this system have now been ordered and assembled. (3) DNA sequence profiling. A DNA profiling technique has been developed that uses automatic fluorescent DNA sequencing to examine hundreds of genomic DNA fragments in a single lane. Alterations in the profile of fragments provide clues about DNA fragments implicated in disease.

(1) PFG fragments as FISH probes

The main goal of this project is to develop a general method that utilizes genomic DNA fragments as a direct source of DNA for mapping and for generation of specific probes to clone regions from human chromosome 20 that are amplified in breast cancer cells.

First, we demonstrated that four different restriction enzymes (*Not* I, *Fse* I, *Asc* I, *Sgr*A I) can be used to generate large restriction fragments from the human genome.

For this, genomic DNA from the mouse hybrid cell line A9NEO containing human chromosome 20, was digested with different restriction enzymes (eight enzymes were tested, *Sgf* I, *Asc* I, *Fse* I, *Pme* I, *Pac* I, *Not* I, *Sgr*A I, *Swa* I), PFGE-fractionated and hybridized with *Alu*-specific probe. As the *Alu*-specific probe we used random primed labeled 2 kb *Hind* III-*Xba* I restriction fragment from plasmid 5X ALU, containing 5 tandem copies of different partially homologous *Alu* sequences.

Hybridization data allowed us to optimize the PFG running condition for better resolution of particular regions of chromosome 20 containing different fragments with similar sizes. Four different PFG conditions were used to resolve different size ranges of *Not* I and *Sgr*A I fragments (i.e. 0.05-0.5 Mb, 0.5-1.0 Mb, 1.0-3.0 Mb, 3.0-7.0 Mb). Individual PFG lanes were sliced into 1.0-1.5 mm pieces using a sterile glass cover slip for each slice. Each gel slice was placed in 100 µl (about twice the volume of the gel slice) of TEA buffer (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 20 mM 2-Aminoethanol). Each sample was melted at 95°C for 15 min and stored at -20°C. Earlier we demonstrated that in the presence of ethanolamine, agarose DNA samples can be melted up to 100 times (5 min each time at 95°C without significantly decreasing yield of PCR products, (the paper in preparation) before taking an aliquot for PCR.

DNA from PFG slices was directly analyzed by using PCR for the presence of single-copy sequences such as STS D20S120 and PTPN1 mapped in two regions on chromosome 20, amplified in many breast cancer cell lines and primary tumors. PCR analysis was done in two steps. First, the subsets of the PFG slices with the most resolved *Not* I fragments were combined into groups (12 slices into each group) and used as a template for primers specific for PCR with STS, D20S120, and gene, PTPN1. Then individual slices from positive groups were analyzed by using PCR with the same pairs of primers. According to the PCR data, PTPN1 and D20S120 are located on *Not* I fragments with the sizes of 0.54 MB and 0.3 MB respectively.

DNA from a *Not* I PFG slice containing the D20S120 $(CA)_n$ microsatellite marker was amplified by "long" inter-*Alu* PCR with primer TC65. To demonstrate that inter-*Alu* PCR products from PFG slices originated from unique human DNAs, Southern hybridization was performed. The PCR products from the slices are strongly homologous to PCR products from the hybrid cell line and do not hybridize with mouse DNA. The specificity of long inter-*Alu* PCR fragments was confirmed by FISH experiments done by Joe Gray. Both nick-translation-labeled and random primed-labeled probes hybridize to 20q13, where the amplicon is mapped.

Some other hybridization experiments are in progress now. The main goal is to estimate the size of amplified region on chromosome 20 in breast cancer cell lines. PFG-fractionated restriction enzyme-generated DNA fragments from different breast cancer cell lines will be hybridized by Southern with labeled probes mapped in two

amplified regions on chromosome 20. Four breast cancer cell lines in which the amplification is relatively strong (BT-474, SK-BR-3, MDA-157, UACC-812) will be used (Tanner *et al*). Pilot hybridization experiments with two breast cancer (BT-474 and SK-BR-3) and normal A9neo20 cell lines show that PTPN1 gene, which maps in the middle of the amplified regions, hybridizes with fragments *Asc* I (0.77 MB), *Sgr*A I (~0.8 MB) and *Not* I (0.54 MB). The same probe hybridize with different *Fse* I fragments (0.58 MB, 0 49 MB and 0 54 MB respectively in the cell lines BT-474 and SK-BR-3 and A9neo20).

(2) cDNA Analysis

A major goal of the initial project was to make dense arrays of cDNAs and use these in two-color comparative hybridization experiments to profile gene expression differences in differential hybridization. A recent report by Dari Shalon and collaborators indicates the feasibility of this proposed approach on a much simpler system, yeast cDNAs. Shalon has developed procedures for printing dense arrays of DNA samples and for two-color imaging (Schena *et al*). He has offered to work with us to make model arrays as we test to see if the method can be extended to much more complex DNA samples. These specific objectives have been defined to optimize two-color cDNA profiling in human samples.

(a) Arrays with better cDNA hybridization properties. Samples used by Shalon and others in the past have consisted of cDNAs baked onto glass. In these samples cDNAs are attached in multiple places, and the overall hybridization kinetics and efficiency are poor. This does not matter for low complexity samples, but it will affect performance for the high complexity samples we expect in breast cancer cells. We have experimented with a number of approaches for improved surface immobilization of DNA samples including 5'-biotinylated DNA bound to streptavidin surfaces and 5'-thiol containing DNAs covalently derivatized to surfaces. We will use these improved surfaces to make the arrays needed for this project. The surfaces show DNA loading of 20 to 200 fmol/mm^2. These approach maximum saturation of the flat surfaces--thus they should give very good signals.

(b) Padlock probe arrays for background reduction. A major problem with hybridization analysis of complex DNA samples is non-specific background. This arises both from cross hybridization and from non-specific absorption of labeled probes. Both problems can be suppressed markedly by a newly developed procedure described by Ulf Landegren called padlock probes. This procedure uses ligation to topologically link a probe and a target. In order for this to occur one of the two must be a single-stranded circle, and in the results shown by Landegren this was the target, rendered topologically circular through multipoint surface attachments. To be able to combine the best surface immobilization with the padlock approach we are testing two approaches. In one, a circular target will be surface immobilized and then used for hybridization with a linear padlock probe that will cyclize after

ligation. In the second approach the ends of an immobilized linear target will be cyclized after hybridization to a circular probe by ligation. For both of these approaches we need to create probes immobilized at internal positions. In experiments thus far, we have successfully reproduced padlock probe reactions in model systems. This gives us considerable confidence that the approach we outline here will be successful.

(c) Subtracted cDNA libraries. Since the original proposal was written, procedures for cDNA subtractions have become substantially improved, most notably through the work of our collaborator on this project, Eugene Sverdlov (Ermolaeva and Sverdlov). This now suggests that instead of straight comparative cDNA profiling we employ subtractive cDNA profiling since it should yield a much simpler pattern of results, one that should be much more compatible with conventional views of diagnostic tests. As part of this project we will implement this new method and compare the breast cancer cell lines in hand in collaboration with Dr. Sverdlov.

After evaluating a number of options for fluorescent imaging of cDNA arrays, we decided to put together a custom-designed Zeiss Axiovert 100 (Carl Zeiss, Germany). The Axiovert 100 was modified (by manufacturer) to be optimized in the epi-fluorescent imaging mode which is needed to image fluorescently-labeled molecules immobilized on an opaque substrate, i.e., silicon.

A 16-bit charged-coupled detector (CCD) camera (Princeton Instruments, NJ) is interfaced to the Axiovert 100. Small changes in fluorescence intensities can be easily detected using this CCD camera. Furthermore, the low light intensities sometimes associated with fluorescence can be overcome by operating the CCD camera in a cumulative mode whereby consecutive images are collected and summed together over an interval of time. Typically, images are collected and summed together until the maximum intensity reaches 90% of the dynamic range. Imaging two different fluorescent colors will employ filter optics to separate the individual fluorescent signals, and image collection is performed twice, i.e., once for each filter set. Images will then be analyzed using the image analysis software, TN-Image (National Institutes of Health).

Robotics interfaced to the Axiovert 100 include an X-Y translation stage (Ludel) to locate and record the positions of interesting sites on the cDNA arrays for further investigations. A robotic shutter (Ludel) allows the control of illumination time so photon-induced bleaching of the fluorochrome can be minimized. All robotics as well as the CCD camera is controlled by a host computer (Gateway 2000, 586-120 Mhz) using the Microsoft Windows 95 operating system. The digital imaging system is mounted on a vibration isolation system to reduce noise associated with long cumulative sampling times. This system has just been

assembled and tested and the first experiments in fluorescent nucleic acids will commence shortly. Experiments will be performed in a class 100 cleanroom to minimize exposure of the cDNA arrays to dust particles.

(3) Profiling DNA Sequence Mixtures

The major thrust of the initial grant proposal was to develop and test two methods of comparative DNA hybridization and then apply these to look for DNA or RNA differences significant in distinguishing characteristics of different breast cancers. Both methods initially conceived are potentially high throughput assays capable of looking at a significant portion of the genome in a single experiment. However, both will require considerable effort to implement in an automated format.

After the grant was submitted I conceived of an alternative method of DNA profiling that, if successful, would be easier to implement because its automation could be carried out without the development of any new procedures. During the past year this new method has been tested, and I am pleased to report that it has performed even better than our original expectations. In the paragraphs below I describe the new method and our initial results with it. Our plan, in the next year of the project, is to apply this method to search for DNA differences involved in breast cancer.

Automated fluorescent DNA sequencers are capable of detecting and quantitating many DNA fragments in a single gel electrophoretic lane. Ordinarily these instruments are used to analyze a nested set of fragments generated from a single DNA species in order to be able to read its sequence. Our approach is to take advantage of this automated equipment, but use it instead to analyze the pattern of DNA species in a very complex mixture. The objective is to compare DNA samples that are as similar as possible except for any differences that may be related to the particular biological target of the investigation. In our work we use the Pharmacia automated laser fluorescence (ALF) instrument. It can analyze 40 samples at once in a run that typically takes six hours. The ALF can resolve a few hundred DNA bands in the size range of 0 to 1 kb. Hence, in order to optimize its use, the target samples of interest must be reduced in complexity to this level.

Our key challenge was to find a way to reduce the complexity of genome DNA quantitatively and reproducibly while directing the focus of the investigation towards a class of target sequences likely to be of biological interest. Since tandemly-repeating DNA sequences have high rates of mutation and are known to be involved in many human diseases, we selected these sequences for our initial studies. The methods developed are quite general, however, and could easily be applied to other classes of interesting DNA sequences. The procedure we use is summarized in Figure 1.

1. DNA was prepared from blood samples in standard ways and cut to completion with one of several different restriction enzymes. Splints of known

sequence were ligated onto the ends of the genomic DNA fragments.

2. DNA was denatured, and specific sequences were captured by hybridization with streptavidin-coated magnetic microbeads containing the complement of the sequence of interest. The resulting immobilized samples, for a typical trinucleotide repeat would contain thousands of fragments, and thus they are too complex to analyze directly by an ALF instrument.

3. Captured DNA was amplified, and a label was introduced by PCR. The primers selected also allowed a controlled reduction of the complexity of the sample. One primer corresponded to the ligated splint, while the second, labeled, primer was complementary at its 3'-end to the trinucleotide repeat but anchored by one or more unique bases at its 5'-end. This novel primer design ensured (a) that the amplified product would contain the entire repeat and (b) that a complexity reduction would be achieved, such that different portions of the initial mixture of target fragments could be selected uniquely and reproducibly just by altering the primer anchor used.

4. The PCR products were denatured and analyzed on the ALF. They showed strikingly complex but quantitatively reproducible profiles of DNA fragments. Several tests of the sensitivity and reproducibility of the entire process have been carried out. Multiple samples from the same individual show a superimposable pattern of DNA bands. Samples from identical twins are usually superimposable also. Samples from simple families show many DNA bands in common as would be expected statistically from simple linkage considerations. However, samples from unrelated individuals are so different and distinct that the method is really worth serious investigation as a potentially powerful forensic tool.

Before attempting to use this new method on breast cancer samples, I decided to test it on two interesting sets of DNA already available in the laboratory. Huntington's Disease (HD) is known to be caused by an expanded triplet repeat, and I asked whether this expanded allele could actually be detected in a complex genome mixture. Although the HD repeat behaves rather badly in PCR and gives fragment sizes that are strongly affected by PCR reaction conditions I was able to see a unique DNA band in complex profiles that correlated with the HD phenotype. The samples used all derived from the Venezuela HD pedigree. In the second test of the method I compared several sets of monozygotic twins discordant for schizophrenia. Indirect evidence had suggested that some cases of schizophrenia might involve triplet repeat expansion. In fact, an identical new DNA band is seen in three of the first four discordant affected twins examined. This is an extraordinarily exciting and quite fortuitous finding, but it does reflect the enormous power of the highly parallel DNA profiling method we have developed.

To use the new method on breast cancer I need to locate samples that are likely to be genetically identical except for DNA sequences that may be altered in

breast cancer. Instabilities in the length of simple repeating sequences have been detected in cells mutant in mismatch repair (Kim et al., Fishel *et al.*, and Parsons *et al.*). Mutations in the mismatch repair system are associated with colon cancers. Two approaches will initially be undertaken. I will compare populations of triplet repeats in the five available breast cancer cell lines, and in cell lines described by Ceriani *et al.* The later samples produce tumors in nude mice. Thus, profiles between the solid tumors and tissue lines will be compared. Eventually, this analysis will be extended to primary tumors.

In the second approach an alternative approach for complexity reduction, potentially relevant for breast cancer will be tested. PFG bands in regions known to be susceptible to amplification in breast cancer have been identified as described in section (1) in this report. These beads are already near the level of sequence complexity where DNA profiling is possible. Bands of interest will be recovered from PFG fractionations, cut with restriction enzymes, and splints will be added by ligation. Direct PCR products from labeled splint-complementary DNA will be used for profiling. If necessary the complexity of this mixture will be reduced by a single-sided *Alu* + splint PCR procedure that we have developed for other purposes. Because the different cell lines available to us come from unrelated individuals, I expect to see many DNA differences. Hence, in this case what will be of interest, originally, is any new bands, not seen in DNA from normal individuals, but seen in more than one of the breast cancer cell lines or tumors derived from the growths of these cells in nude mice.

Conclusions.

The DNA sequence profiling method outlined in the report is extremely promising because it has, potentially, a hundred times the throughput of methods currently in use. To successfully implement it in the study of breast cancer will require several developments including the selection of appropriate target DNA samples, and the development of automated software to analyze the very complex data patterns that result. cDNA profiling has been shown by others to be practical in simple organisms and we are refining the method to produce the highest possible data quality in complex organisms. A particularly attractive notion which will form the basis for our future work is combining this method with cDNA subtraction techniques to simplify the process of data collection and analysis and to highlight potentially significant differences in advance. Analysis of genomic DNA differences in breast cancer cell lines confirms the importance of chromosome 2 regional amplification in breast cancer. Now, what needs to be done is to make a fine map of these regions. This is a procedure with which we have had much experience and it should be relatively easy to accomplish in the next year.

References.

Ceriani, R. L., Peterson, J. A., Blank, E. W., Chan, C. M., and Cailleau, R. 1992. Development and Characterization of Breast Carcinoma Cell Lines as *In Vitro* and *In Vivo* Models for Breast Cancer Diagnosis and Therapy. In Vitro Cell. Dev. Biol. 28A: 397-402.

Ermolaeva, O. D., Sverdlov, E. D. 1995. Subtractive Hybridization, a Technique for Extraction of DNA Sequences Distinguishing Two Closely Related Genomes. Genetic Analysis (Biomolecular Engineering), in press.

Fishel, R., Lescoe, M. K., Rao, M. R., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M., and Kolodner, R. 1994. The Human Mutator Gene Homolog MSH2 and its Association with Hereditary Nonpolyposis Colon Cancer. Cell 77: 167.

Kim, H., Jen J., Vogelstein, B., and Hamilton, S. R. 1994. Clinial and Pathological Characteristics of Sporadic Colorectal Carcinomas with DNA Replication Errors in Microsattelite Sequences. Am. J. Pathol. 145: 148-156.

Parsons, R., Li G. M., Longley, M., Modrich, P., Liu, B., Berk, T., Hamilton, S. R., Kinzler, K. W., and Vogelstein, B. 1995. Mismatch Repair Deficiency in Phenotypically Normal Human Cells. Science 268: 738-730.

Schena, M., Shalon, D., Davis, R. A., and Brown, P. O. 1995. Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. Science 270: 467-470.

Tanner, M., M. Tirkkonen, A. Kallioniemi, C. Collins, T. Stokke, R. Karhu, D. Kowbel, F. Shadravan, M. Hintz, W.-L. Kuo, F. M. Waldman, J. J. Isola, and J. W. Gray. 1994. Increased Copy Number at 20q13 in Breast Cancer: Defining the Critical Region and Exclusion of Candidate Genes. Cancer Research 54: 4257-4260.

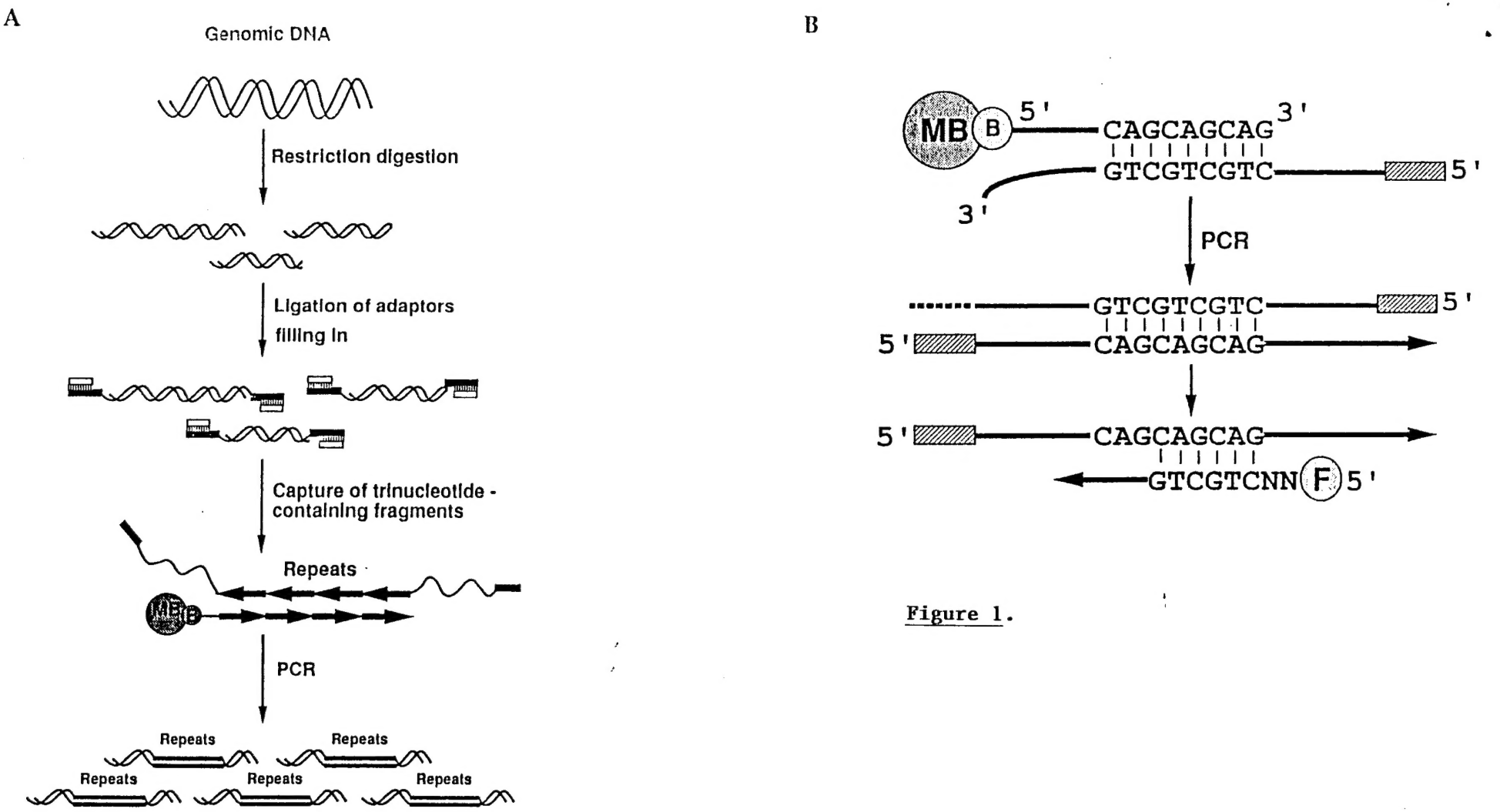

Figure 1.

Figure 1. Capture (A) and Labeling (B) of genomic restriction containing an interspersed trinucleotide repeating sequence. Genomic DNA is cleaved by a restriction enzyme and end-tagged with known sequences using DNA ligase. The restriction fragments are incubated with an immobilized capture probe containing a sequence complementary to the interspersed repeat sequence. The captured fragments are removed from the probe sequence and fluorescently labeled using an anchored (at the 5' end) DNA polymerase primer at the 5'-end of the interspersed repeat to ensure the labeled molecules contain the full length of the repeat.